

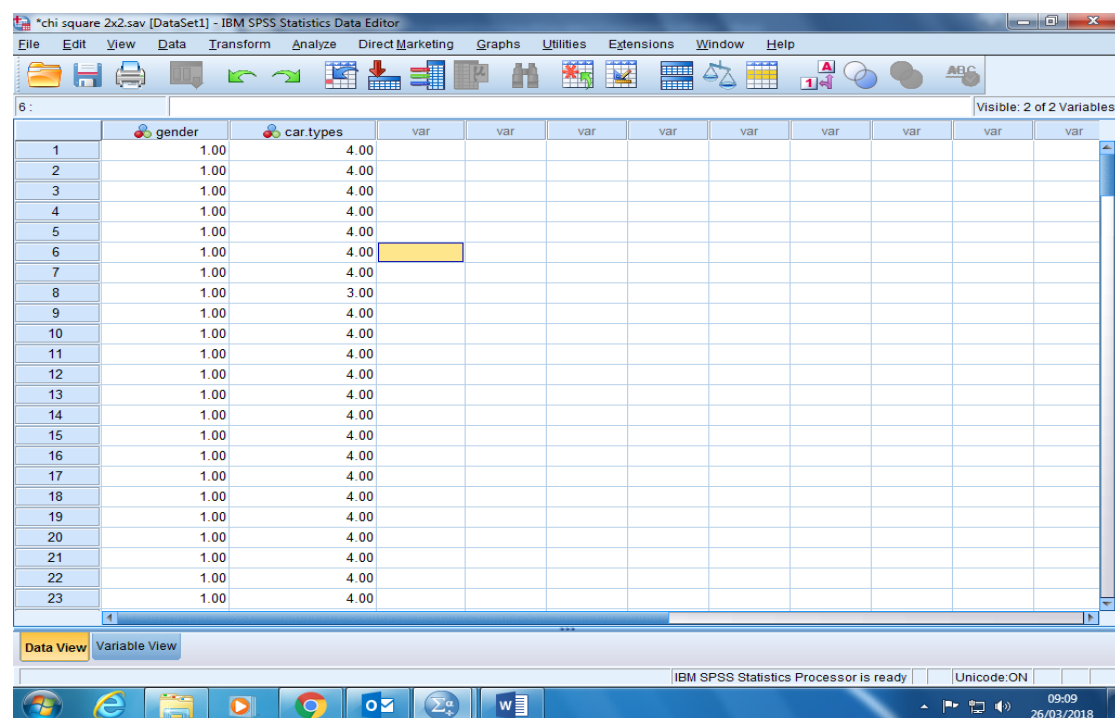
Research Methods

Two variable Chi-squared test on SPSS

We will use SPSS to do a chi-squared test on the two-way table used in the lecture (the *gender x type of car* cross-tabulation). The table is reproduced below:

	Ford	Vauxhall	Mazda	BMW
Male	8	23	31	61
Female	22	31	19	5

Firstly, this table has to be represented in the SPSS data editor. The manner in which you will do this is similar to the way the data file was structured for the unrelated t-test, except that here, there are two between-subjects variables (*gender* and *ford/vauxhall/mazda/BMW*). Therefore, you will need **two coding variables**. So in the data editor, define two variables: *gender* (Male and female) and *cartype* (*ford/vauxhall/mazda/BMW*). The final data file will look as follows:

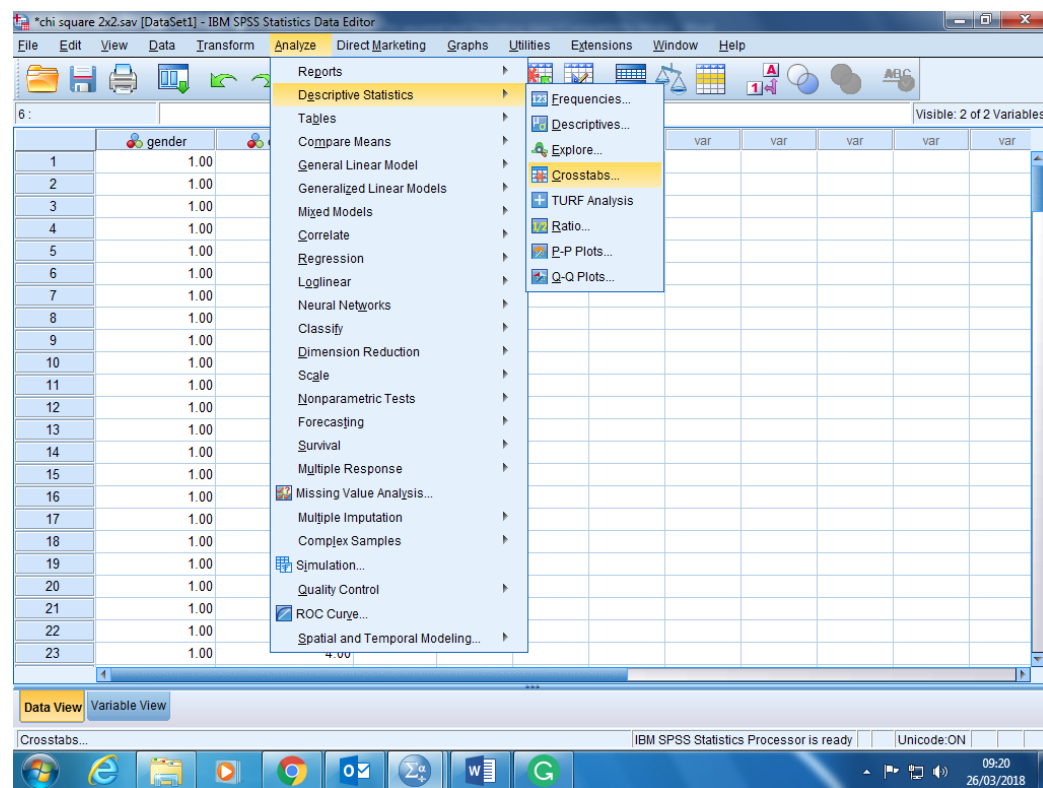


So the first row represents the 200 participants: 123 males and 77 females. In the second row, we have car type: 30 ford, 54 vauxhall, 50 mazda and 66 BMW.

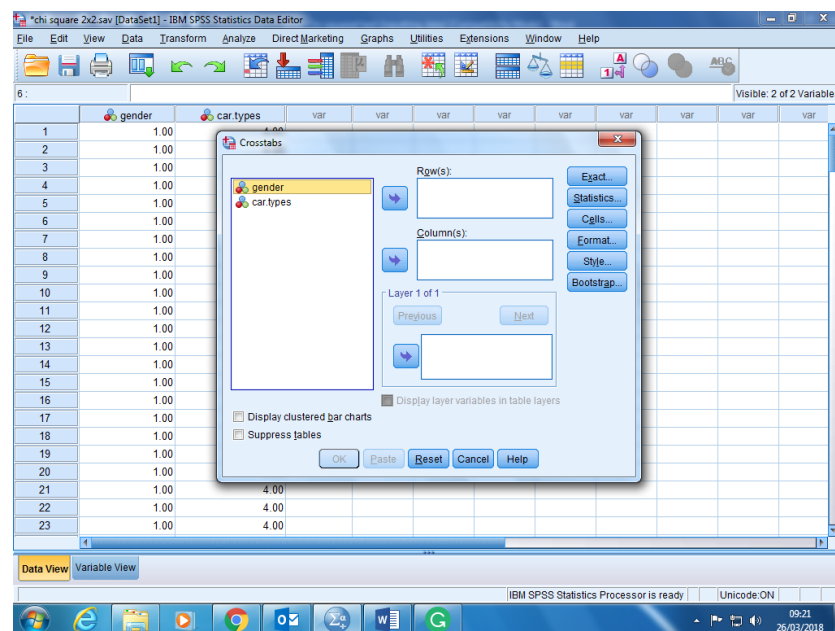
You have already labelled variables in the value boxes – follow this process here.

Do the chi-square analysis

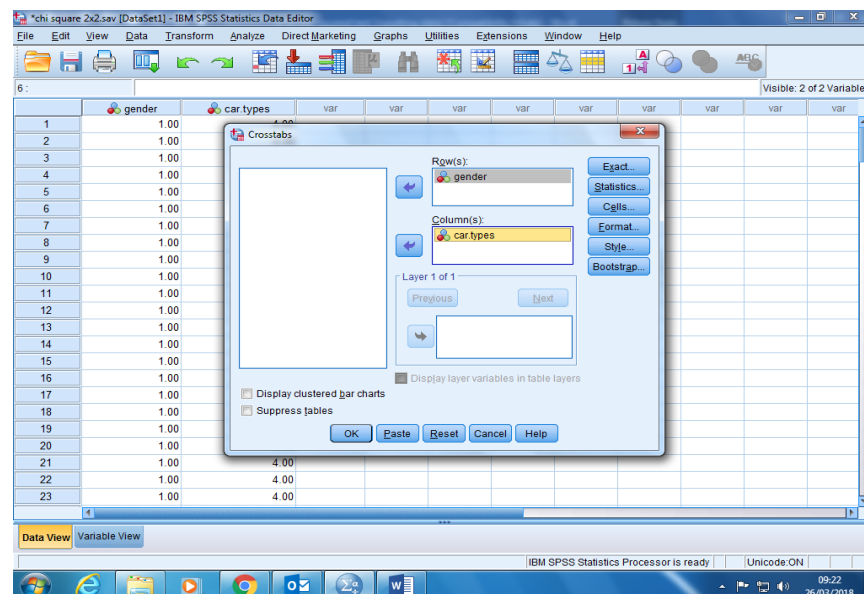
Go to the '**Analyse**' menu. From this select '**Descriptive Statistics**' and from this sub-menu, select '**Crosstabs...**'.



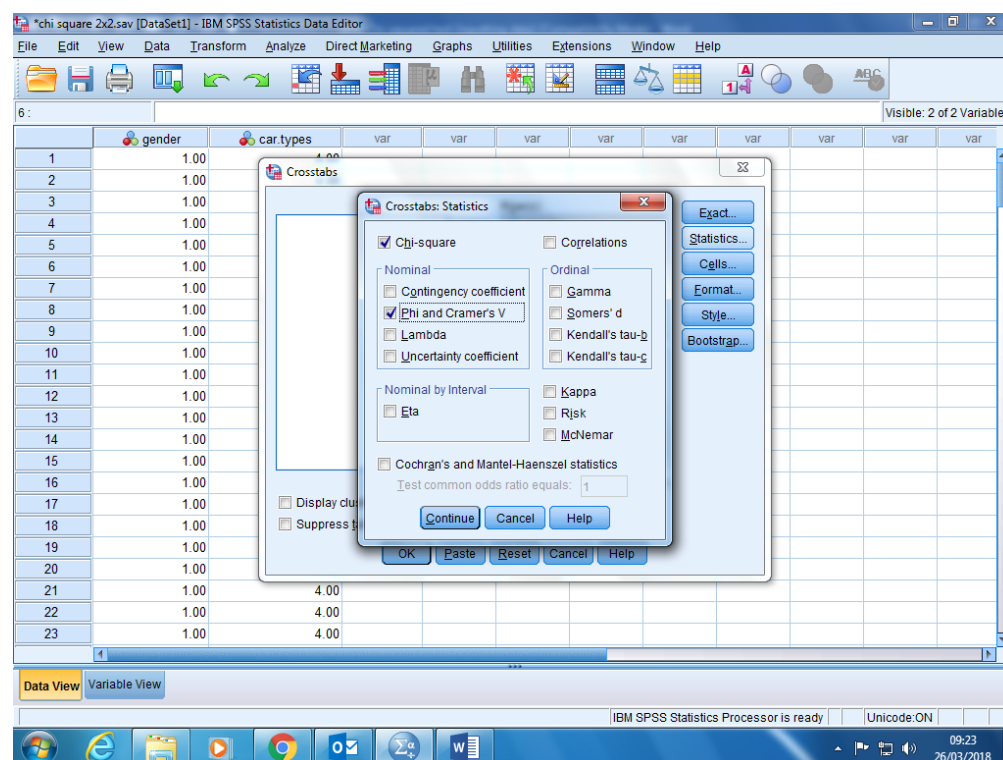
You will obtain the following dialogue box:



Select 'Gender' and place this in the '**Rows**' box by clicking the arrow button next to it. Similarly, select 'Cartype' and place this in the '**Columns**' box:

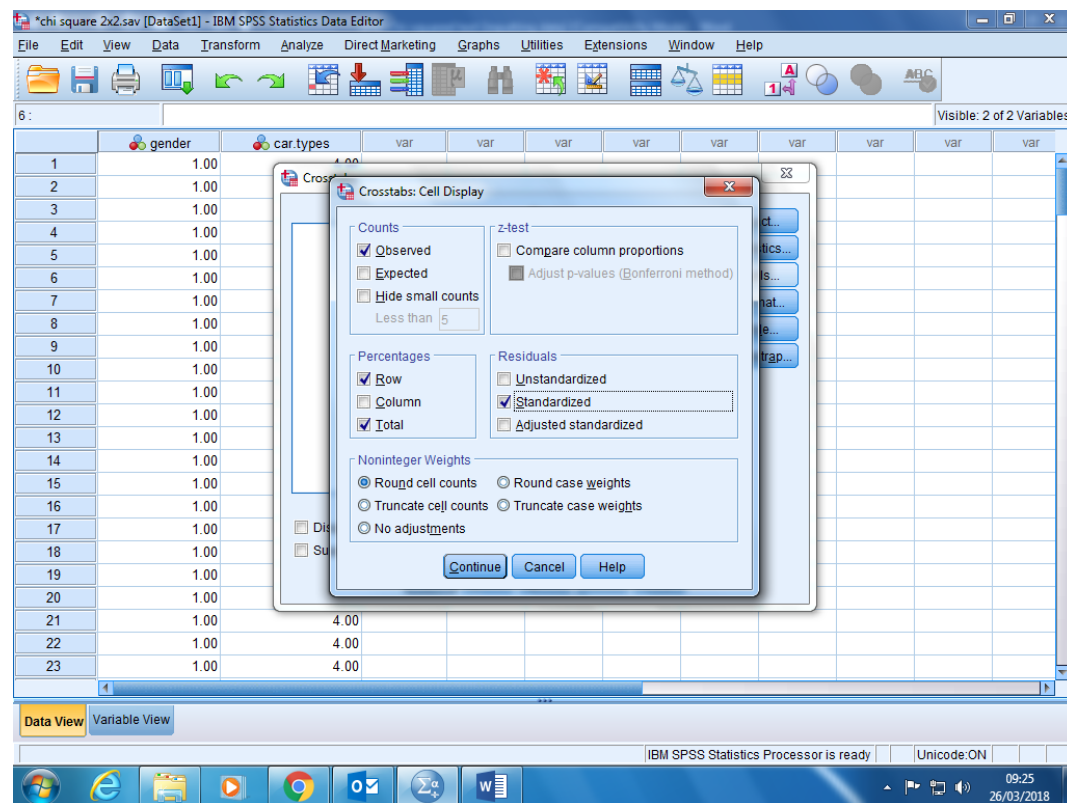


Next, click the '**Statistics**' button:



In the resulting dialogue box, click the '**Chi Square**' box at the top left hand corner of this and the Phi and Cramer's V (for effect size). Click '**Continue**' which will return you to the main dialogue box. Click on Cells and select percentages – row and total and when you have more than 2 categories (you have 4 car-types) you need to

select standardised residuals (this tells us something about which category plays the largest role in the association. Then click continue and **OK**. SPSS will then do the analysis and provide the following output:



After a 'case processing summary' SPSS reproduces the contingency table on which you did the analysis:

gender * car.types Crosstabulation

			car.types				
			ford	vauxhall	mazda	BMW	Total
gender	male	Count	8	23	31	61	123
		% within gender	6.5%	18.7%	25.2%	49.6%	100.0%
		% of Total	4.0%	11.5%	15.5%	30.5%	61.5%
		Standardized Residual	-2.4	-1.8	.0	3.2	
	female	Count	22	31	19	5	77
		% within gender	28.6%	40.3%	24.7%	6.5%	100.0%
		% of Total	11.0%	15.5%	9.5%	2.5%	38.5%
		Standardized Residual	3.1	2.2	-.1	-4.0	
Total	Count	30	54	50	66	200	
	% within gender	15.0%	27.0%	25.0%	33.0%	100.0%	
	% of Total	15.0%	27.0%	25.0%	33.0%	100.0%	

The next table provides the results of the Chi-Squared test:

Chi-Square Tests

	Value	df	Asymptotic Significance (2- sided)
Pearson Chi-Square	50.189 ^a	3	.000
Likelihood Ratio	56.298	3	.000
Linear-by-Linear Association	48.819	1	.000
N of Valid Cases	200		

a. 0 cells (0.0%) have expected count less than 5. The minimum expected count is 11.55.

The next table provides the results of the Cramer's V (used when more than 2 categories):

Symmetric Measures

	Value	Approximate Significance
Nominal by Nominal Phi	.501	.000
Cramer's V	.501	.000
N of Valid Cases	200	

You can see that you have a range of test statistics for chi-squared. The first of these – Pearson is the 'standard' way of calculating the statistic and in general this is the one you will use most of the time. You can see that the resulting value of χ^2 of 50.189 has a significance value of .000, indicating a statistically significant result.

To report this result:

Differences in car-type choice between different gender were apparent – the majority of males were more likely to choose BMW, whereas the majority of females were more likely to choose Ford. A chi-squared test indicated a significant association between car-type choice and gender ($\chi^2 = 50.189$, $df = 4$, $p < .05$). Cramer's V shows that the effect size of the association was large (.501)

50% males – chose BMW compared to only 7% of females. 29% females chose Ford compared to 7% of males

Important note about Chi-squared Statistic

As mentioned above, most of the time we would use the **Pearson** value as our chi-squared statistic. The one exception to this is cases where we have a 2x2 cross tabulation (i.e. two rows and two columns). If you do a chi-squared test on this type of table, you will see that in the results you get an additional value of chi-squared called a '**continuity correction**'. In the case of a 2x2 table, this provides the most accurate value for the test statistic, so this is the value you should use, the interpretation is the same as any other test statistic: The significance value should be less than or equal to .05.